

1. はじめに

地域をサンプルとする空間データには、地域と地域が相互依存する空間的依存性や、集計単位によってデータの示す傾向が異なるといった性質がある。後者の性質は MAUP¹⁾と呼ばれ、前者の性質とあわせて、ゾーンのスケールや形状等の地理的な単位の設定が分析結果に影響を及ぼすことをあらわす。同一の空間計量経済モデルを集計単位の異なるデータに適用したとき、MAUP の影響によって、空間相関パラメータや、構造パラメータの大きさがばらついて、統計的に有意な結果が得られない、または期待される符号が得られないなどの影響を及ぼす。

本研究では、空間計量経済モデルのパラメータが安定的に推計されない問題の原因を、詳細な地域データでは、地域間のばらつきの大きさが強調される傾向にあるため、と考える。そこで、面補間法の考えを援用して、集計後の大ゾーンを単位とする空間集計行列を用いた、空間計量経済モデルを開発する。また、推定されたモデルを用いてパラメータ推計値の安定性と、小ゾーン予測性能の検討を行う。

2. 空間集計モデル

1) 体積保存則と空間集計行列

本研究では、大ゾーン（地域数 M ）でのみ目的変数 $\mathbf{Y}_L$ が得られ、小ゾーン（地域数 N , $N > M$ ）では説明変数 $\mathbf{X}_S$ しか得ることができない空間を想定する。なお、小ゾーン単位のデータは添え字 S を、大ゾーン単位のデータは添え字 L を用いる。

空間計量経済モデルが有すべき性質に、体積保存則がある。これは、小ゾーン単位の予測値を集計して得られる大ゾーンの予測値が、実際の大ゾーンの観測値に等しい性質であり、式(1)のように表す。

$$y_l^L = \sum_j^N b_{lj} \hat{y}_{iej}^S \quad (1)$$

y_l^L : 大ゾーン l の観測値, $\hat{y}_{iej}^S$: 大ゾーン l に含まれる小ゾーンの予測値, b_{lj} : 空間集計行列の要素。

空間集計行列は行側に大ゾーン、列側に小ゾーンを

とり、小ゾーンが大ゾーンに帰属する場合は 1、それ以外は 0 となる、 M 行 N 列の行列である ($N > M$)。

空間集計行列 $\mathbf{B}$ を、式(2)で定義する。

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & 1 \end{bmatrix} \quad (2)$$

空間集計行列の要素 b_{lj} は、ゾーン集計後の大ゾーン l に含まれる小ゾーン数を m_l としたとき、 $\sum_{i=1}^M b_{ij} = 1$, $\sum_{j=1}^N b_{lj} = m_l$ を満たすように定義されている。

2) モデルの定式化

現在、空間計量経済モデルとして、目的変数に空間的自己相関を考慮した空間自己回帰モデル (SAR : Spatial Auto-Regressive) や誤差項に空間的自己相関を考慮した空間移動平均誤差モデル (SMA : Spatial Moving Average), およびそれらを複合した空間自己回帰移動平均モデル (SARMA) などが知られている。Manski は、上述の一連のモデルを特殊系とする、より一般的な Manski モデルを提案している²⁾。Manski モデルを式(3)に示す。

$$\begin{cases} \mathbf{Y} = \rho \mathbf{W} \mathbf{Y} + \mathbf{X} \boldsymbol{\beta} + \mu \mathbf{W} \mathbf{X} \boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} = \lambda \mathbf{W} \boldsymbol{\varepsilon} + \boldsymbol{\varepsilon} \end{cases} \quad (3)$$

ただし ρ , λ , μ : 空間相関パラメータ, $\mathbf{W}$: 空間近接行列, $\boldsymbol{\beta}$: 構造パラメータ, $\boldsymbol{\varepsilon}$: 誤差項である。

本研究で提案する空間集計モデルとは、式(3)に左から空間集計行列 $\mathbf{B}$ を乗じて、地域数 N から地域数 M へ地域を集約したモデルである。定数項を加え、整理すると式(4)のようになる。

$$\begin{aligned} \mathbf{Y}_L &= \mathbf{B} \mathbf{Y}_S \\ &= \mathbf{B} (\mathbf{I}_N - \rho \mathbf{W}_S)^{-1} (\mathbf{I}_N + \mu \mathbf{W}_S) \mathbf{X}_S \boldsymbol{\beta} \\ &\quad + \mathbf{B} (\mathbf{I}_N - \rho \mathbf{W}_S)^{-1} \boldsymbol{\gamma} + \mathbf{B} (\mathbf{I}_N - \rho \mathbf{W}_S)^{-1} (\mathbf{I}_N + \lambda \mathbf{W}_S) \boldsymbol{\varepsilon}_S \end{aligned} \quad (4)$$

$\mathbf{I}_N$: N 行 N 列の単位行列, $\boldsymbol{\gamma}$: 定数項。ここで、

$$\begin{cases} \mathbf{A} = \mathbf{I}_N - \rho \mathbf{W}_S \\ \mathbf{F} = \mathbf{I}_N + \lambda \mathbf{W}_S \\ \mathbf{G} = \mathbf{I}_N + \mu \mathbf{W}_S \end{cases} \quad (5)$$

とおくと、式(4)は式(6)のようにさらに整理できる。

$$\begin{aligned} Y_L &= BY_S \\ &= BA^{-1}GX_S\beta + BA^{-1}\gamma + BA^{-1}F\epsilon_S \end{aligned} \quad (6)$$

3) モデル推計手順

式(4)の各パラメータは最尤法によって推計される。最尤法推計に用いる対数尤度関数を式(7)に示す。

$$\begin{aligned} \ln L(\xi) &= -\frac{M}{2} \ln(2\pi) - \frac{M}{2} \ln(\sigma^2) \\ &\quad - \frac{1}{2} \ln \left| (BA^{-1}F)(BA^{-1}F)' \right| \\ &\quad - \frac{1}{2\sigma^2} (F^{-1}AB^{-1}BA^{-1}F\epsilon_S)' (F^{-1}AB^{-1}BA^{-1}F\epsilon_S) \end{aligned} \quad (7)$$

式(7)には空間分配行列 B^{-1} が現れる。 B^{-1} は、大ゾーン単位の目的変数 Y_L を小ゾーン単位の分配する行列であり、体積保存則を満たすように定義する。空間分配行列の要素 b^{ij} を式(8)のように定式化する。

$$b^{ij} = \frac{\hat{y}_{i \in I}^S(\rho, \lambda, \mu, \beta)}{\sum_{i=1}^{m_j} \hat{y}_{i \in I}^S(\rho, \lambda, \mu, \beta)} \quad (8)$$

小ゾーンの予測値 $\hat{Y}_s$ は、空間相関パラメータ $\hat{\rho}$, $\hat{\lambda}$, $\hat{\mu}$ と構造パラメータ $\hat{\beta}$ の関数として定式化した空間分配行列 $B^{-1}(\hat{\rho}, \hat{\lambda}, \hat{\mu}, \hat{\beta})$ を用いて、式(9)によって求める。

$$\hat{Y}_s = B^{-1}(\hat{\rho}, \hat{\lambda}, \hat{\mu}, \hat{\beta}) Y_L \quad (9)$$

3. 幹線旅客純流動データを用いた実証分析

1) データの概要

本研究では、小地域補間の可能性を探るため、詳細地域での観測が困難な集中交通量データを分析対象として用いる。なお、提案モデルの性質を明らかにするため、小地域補間について真値が得られるデータ集計単位での検討を行う。具体的には、全国の生活圏 207 地域のうち、沖縄と離島を除く 194 地域を小ゾーン S 、沖縄を除いた 46 都道府県を大ゾーン L とし、提案モデルの小地域補間への適用可能性を検討する。ここで、目的変数は集中交通量、説明変数は民営事業所従業者数、宿泊業・飲食店数、一般化費用とする。また、集中交通量データとして、第 4 回全国幹線旅客純流動調査データ (2005 年) を用いる。

2) モデル推計結果

式(6)は空間相関パラメータ ρ , λ , μ を除く (0 に固定する) ことで様々な型に変形できる。ここでは、

最終尤度が最も高かった SARMA 型 ($\mu = 0$) について考察を行う。SARMA 型の空間集計モデルを式(10)に示す。

$$\begin{aligned} Y_L &= BY_S \\ &= BA^{-1}X_S\beta + BA^{-1}\gamma + BA^{-1}F\epsilon_S \end{aligned} \quad (10)$$

また、小ゾーンのデータに基づいて推計されるモデル (以下、小ゾーンモデル) についても推計を行い、空間集計モデルの推計結果と比較を行う。SARMA 型の小ゾーンモデルを式(11)に示す。

$$Y_s = A^{-1}X_s\beta + A^{-1}\gamma + A^{-1}F\epsilon_s \quad (11)$$

表 1 に空間集計モデル、小ゾーンモデルの推計結果 (SARMA 型) をそれぞれ示す。以下で決定係数とはモデルの推計に用いたゾーンスケールに対する値である。

表 1 モデル推計結果 (SARMA 型)

	空間集計 モデル	小ゾーン モデル
変数	推計値	推計値
ρ (空間相関パラメータ)	-0.375	-0.199 *
λ (空間相関パラメータ)	0.490 +	0.762 **
β_1 (民営事業所従業者数)	0.022 **	0.025 **
β_2 (宿泊業・飲食店数)	-0.155	-0.202 **
β_3 (一般化費用)	-2.842 **	-2.242 +
γ (定数項)	4859.001 **	3427.056 *
σ (誤差の標準偏差)	3967.380 **	2900.427 **
M Moran's I 検定統計量	8.240	0.958
R^2 決定係数	0.941	0.892
$RMSE$	2246.801	2955.818
最終尤度	-2.425	-9.367
$\hat{y}_s < 0$ の数	5	4
サンプル数	46	194

+ 10%有意 * 5%有意 ** 1%有意

小ゾーンモデルにおいて有意となった目的変数の空間相関パラメータ ρ は、提案モデルでは有意とならなかった。しかし、 t 値は 1.63 となり、10%有意水準の値である 1.64 に非常に近い値であった。また、説明変数 X_2 (宿泊業・飲食店数) のパラメータ β_2 も提案モデルでは有意とならなかった。これは変数間に極めて高い重共線性があったために発生した問題である。

さらに、小ゾーンモデルでは残差における空間的自己相関が統計的には存在しなかったのに対し、提案モデルでは残差における空間的自己相関が存在するという結果となった。その一方で、最終尤度をはじめとする、モデルの推計精度をあらわす各指標は、提案モデルのほうが、小ゾーンモデルよりも優れていた。

3) 小地域補間への適用可能性

提案モデルによる小ゾーン単位の予測値 $\hat{Y}_s$ は式(9)より算出する. 小ゾーン単位の予測値 $\hat{Y}_s$ と実際の観測値 Y_s の関係を図 1 に示す.

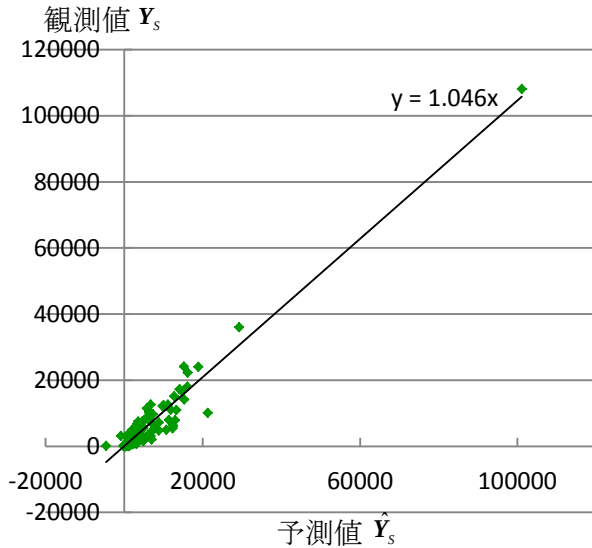


図 1 空間集計モデルの予測値と観測値の回帰直線

予測値 $\hat{Y}_s$ と観測値 Y_s に対する回帰直線の傾きは 1.046, 相関係数は 0.970 であり, 空間集計モデルの再現性は高い. しかし, 予測値が負の値となる地域が 5 地点発生した. これらは, 集中交通量が 200 以下の地域であり, 集中交通量の少ないゾーンの再現性が低い傾向がある. そこで, 観測値において中央値 1595 よりも大きいゾーンを A1 (97 ゾーン), 小さいゾーンを A2 (97 ゾーン) と区別して, さきほど同様に, 式(9)から推計した小ゾーン単位の予測値 $\hat{Y}_s^{A1}$, $\hat{Y}_s^{A2}$ と観測値 Y_s に対する回帰直線をそれぞれ求めた. 結果を図 2, 図 3 に示す.

図 2 より, A1 の予測値 $\hat{Y}_s^{A1}$ と観測値 Y_s に対する回帰直線の傾きは 1.053, 相関係数は 0.968 であり, 再現性は高い. また, 図 3 より, A2 の予測値 $\hat{Y}_s^{A2}$ と観測値 Y_s に対する回帰直線の傾きは 0.417, 相関係数は 0.534 であり, 再現性は低い. また, 傾きが 0.417 となることから, 予測値は観測値よりも大きな値を示す傾向があることがわかる.

以上の結果から, 提案した空間集計モデルから得られる予測値が大きいゾーンは再現性が高く, 予測値が小さいゾーンは再現性が低い傾向にあることが明らかになった.

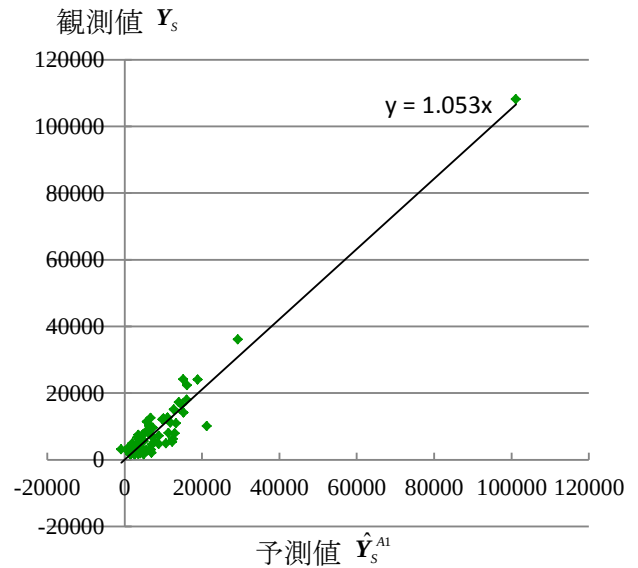


図 2 空間集計モデルの予測値と観測値の回帰直線 (中央値以上のゾーン A1)

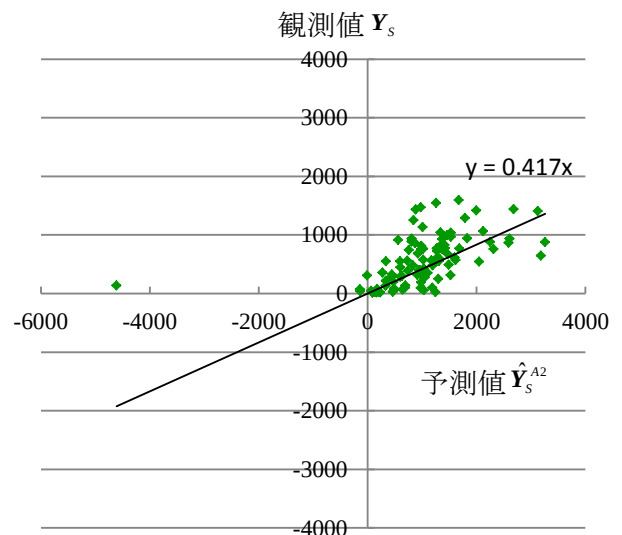


図 3 空間集計モデルの予測値と観測値の回帰直線 (中央値以下のゾーン A2)

4) 異なる集計スケールでの小ゾーン予測値の比較

都道府県単位をさらに集計した 8 地方区分単位を集計大ゾーン $L_{\max}$ (ゾーン数 $M_{\max} = 8$) として, 異なる集計スケールから小ゾーンの予測を行った際の提案モデルの推計精度に関する考察を行う. 表 1 の大ゾーン L の観測値から小ゾーン S の予測値を推計したモデル (以下, モデル[46|194]) の推計結果, 集計大ゾーン $L_{\max}$ の観測値から小ゾーン S の予測値を推計したモデル (以下, モデル[8|194]) の推計結果, 集計大ゾー

表2 異なる単位空間スケールデータに対する推計結果の比較

変数	モデル[46 194]	モデル[8 194]	モデル[8 46]	モデル[(8 46) 194]
	推計値	推計値 ##	推計値	推計値
ρ (空間相関パラメータ)	-0.375	0.989	0.248	-0.375
λ (空間相関パラメータ)	0.490 +	1.000	-0.097	0.490 +
β_1 (民営事業所従業者)	0.022 **	0.015	0.060	0.022 **
β_2 (宿泊業・飲食店)	-0.155	-0.200	-0.835	-0.155
β_3 (一般化費用)	-2.842 **	2.576	0.175	-2.842 **
γ (定数項)	4859.001 **	-3026.683	1741.154	4859.001 **
σ (小ゾーン誤差の標準偏差)	3967.380 **	58.708	8009.205 **	3967.380 **
M Moran's I検定統計量	8.240	——	4.341	0.769
R^2 決定係数	0.941	——	0.877	0.886
RMSE	2246.801	——	9258.478	3871.016
最終尤度	-2.425	——	-1.931	-2.425
$\hat{y}_s < 0$ の数	5	——	1	10
サンプル数	46	8	8	46

+ 10%有意 * 5%有意 ** 1%有意

パラメータが収束していない

ン $L_{\max}$ の観測値から大ゾーン L の予測値を推計したモデル(以下、モデル[8|46])の推計結果、モデル[8|46]による大ゾーン L の予測値 $\hat{Y}[8|46]$ から小ゾーン S の予測値を推計したモデル(以下、モデル[(8|46)|194])の推計結果を表2に、それぞれ示す。

モデル[8|194]では、パラメータが非定常域に発散し、収束しないまま計算が途中で停止したため、推計結果が得られなかった。

モデル[8|46]では、小ゾーン誤差の標準偏差以外のすべてのパラメータが有意とならなかった。さらに、空間相関パラメータ ρ 、 λ と説明変数 X_3 (一般化費用)のパラメータ β_3 の符号が期待される条件を満たさないなど、モデル[46|194]とは異なる結果が得られた。モデル[8|46]では、モデル[46|194]などで用いた空間近接行列ではなく、大ゾーン(都道府県単位)の空間近接行列(M 行 M 列)を用いて推計を行ったためと考えられる。また、Moran's I検定統計量 M は有意であり、残差に正の空間的自己相関が存在する結果となった。一方で、決定係数は高い値を示しており、当てはまりは良好であり、再現性も高かった。

モデル[(8|46)|194]では、目的変数の空間相関パラメータ ρ と説明変数 X_2 (宿泊業・飲食店)のパラメータ β_2 が有意とならなかったものの、それ以外のパラメータは有意となった。また、決定係数は高い値を示しており、当てはまりは良好である。しかし、小ゾーンの予測値が負となる地域が10地点存在した。

以上の推計結果から、モデル[8|194]のように、入手できるデータと、分析者が必要なデータのスケール(ゾ

ーン数)の乖離が大きいことに起因し、推計が行えない場合、モデル[(8|46)|194]のように、それらの中間スケールのゾーンを設定し、段階的に推計を行うことでパラメータは収束し、中間スケールのゾーンデータに対して直接モデル推計した場合と同様の結果が得られることがわかった。

4. おわりに

本研究では、入手できるデータと、分析者が必要なデータのスケールが一致しない場合、後者のスケールにあわせて前者のデータを変換する、面補間法の考え方を空間計量経済モデルに援用して、空間相関パラメータや、構造パラメータの推計結果の安定化を図る手法について検討した。その結果、体積保存則を満たすように定義した空間集計モデルでは、精度の高い小地域補間が可能であることが明らかとなった。

ただし、ゾーンサイズと無関係な、安定した構造パラメータの推計には至っていない。今後、空間的依存性を適切に表現する柔軟なモデル構造を有するモデルを用いて、提案手法の有効性を検証する必要がある。

参考文献

- 1) Openshaw, S. : The modifiable area unit problem, *Geobooks*, 1984.
- 2) Elhorst, J.P. : Applied spatial econometrics: Raising the bar, *SpatialEconomic Analysis*, Vol.5, No.1, pp.9-28, 2010a.